TECHNICAL TRANSACTIONS | CZASOPISMO TECHNICZNE
MECHANICS | MECHANIKA
1-M/2013

RAFAŁ PALEJ*

CUBICALLY CONVERGENT METHOD FOR SOLVING A STANDARD BOUNDARY VALUE PROBLEM

SZEŚCIENNIE ZBIEŻNA METODA ROZWIĄZYWANIA STANDARDOWEGO ZAGADNIENIA BRZEGOWEGO

A b s t r a c t

The paper presents a cubically convergent method for solving a standard boundary value problem consisting of n coupled first-order differential equations and n boundary conditions. The idea of the presented method is based on the shooting method using the expansion of the desired function into Taylor's series including second-order derivatives. Effective use of the iteration formula requires introduction of sensitivity functions and their derivatives. In each iteration, the initial problem, composed of $n(1+n_1+n_1^2)$ first-order differential equations, must be solved, where n_1 signifies the number of unknown parameters. The convergence of the presented method has been illustrated on an example.

Keywords: *shooting method, sensitivity functions, derivatives of sensitivity functions, cubic convergence*

S t r e s z c z e n i e

W artykule przestawiono sześciennie zbieżną metodę rozwiązywania standardowego zagadnienia brzegowego składającego się z n sprzężonych równań różniczkowych pierwszego rzędu i n warunków brzegowych. Idea prezentowanej metody oparta jest na metodzie strzałów wykorzystującej rozwinięcie poszukiwanych funkcji w szereg Taylora uwzględniający pochodne drugiego rzędu. Efektywne skorzystanie ze wzoru iteracyjnego wymaga wprowadzenia funkcji wrażliwości i ich pochodnych. W każdej iteracji należy rozwiązać zagadnienie początkowe składające się z $n(1+n_1+n_1^2)$ równań różniczkowych pierwszego rzędu, gdzie n_1 oznacza liczbę nieznanych parametrów. Zbieżność prezentowanej metody zilustrowano na przykładzie.

Słowa kluczowe: *metoda strzałów, funkcje wrażliwości, pochodne funkcji wrażliwości, zbieżność sześcienna*

* Prof. Rafał Palej, Institute of Applied Informatics, Faculty of Mechanical Engineering, Cracow University of Technology.

1. Introduction

A standard two-point boundary value problem consists of n coupled first-order ordinary differential equations as well as n_1 boundary conditions at one end of the domain and a remaining set of boundary conditions at the other end. There are two classes of numerical methods for solving such a problem: the shooting method and the relaxation methods [1]. The shooting method usually implements a quadratically convergent Newton-Raphson method. Faster convergence can be obtained by taking into account the second-order derivatives in the Taylor's expansion. Proceeding in this way, however, leads to obtaining a set of non-linear equations due to the unknown parameters. The non-linear equations may be "made linear" with the use of the Newton-Raphson solution while the cubic convergence of the iteration process is retained. This approach may be used for iterational solving the set of non--linear algebraic equations [2]. In order to use the iteration formula effectively, we need to introduce not only the sensitivity functions [3] but their derivatives as well. The description of the presented method depends on the manner of the formulation the boundary conditions. The paper presents a situation when each desired function has an imposed boundary condition only at one end of the domain.

2. Formulation of the problem

We shall analyze a boundary value problem defined by a set of n first-order differential equations in the following form:

$$\frac{dy_i}{dx} = f_i(x, y_1, \ldots, y_n), \quad i = 1, \ldots, n \tag{1}$$

For the sake of simplification of the subsequent notation, let us assume that each function $y_i(x)$ has the imposed boundary condition only at one end of the domain. We shall also assume that the number of conditions imposed at the initial point of the domain is at least equal to the number of conditions at the final point. Let us number the desired functions in such a way that the lowest indicators are ascribed to the functions that have the imposed boundary conditions at the final point, i.e.

$$y_i(b) = y_{ib}, \qquad i = 1, \ldots, n_1 \tag{2}$$

and then the functions with indicators greater than n_1 will have the imposed boundary conditions at the initial point

$$y_i(a) = y_{ia}, \qquad i = n_1 + 1, \ldots, n \tag{3}$$

In order to replace the boundary value problem with the initial value problem, we need to complete conditions (3) by introducing n_1 unknown parameters.

$$y_i(a) = p, \qquad i = 1, \ldots, n_1 \tag{4}$$

The desired solutions will now be the functions of the variable x and parameters p_i, i.e.

$$y_i = y_i(x, p_1, \ldots, p_{n_i}), \quad i = 1, \ldots, n \tag{5}$$

We want to find such values of parameters p_i that would enable the initial value problem to fulfil the boundary conditions with the prescribed accuracy (2).

3. The method of finding the solution

We shall expand the functions $y_i = 1, \ldots, n_1$ into the Taylor's series around the trial values of parameters $p_i^{(1)},\ i = 1,\ldots,n_1$ for $x = b$, including the expressions containing second derivatives. The connection between the value of the i-th function at point $\mathbf{p}^{(2)}$ and the values of the said function and its derivatives at point $\mathbf{p}^{(1)}$ – signifying the first approximation of the unknown parameters, is provided by the Taylor's formula

$$y_i(b,\mathbf{p}^{(2)}) = y_i(b,\mathbf{p}^{(1)}) + \nabla y_i^T(b,\mathbf{p}^{(1)})\mathbf{h}^{(1)} + \frac{1}{2}\mathbf{h}^{(1)T}\mathbf{H}_i(b,\mathbf{p}^{(1)})\mathbf{h}^{(1)} + \ldots,\quad i = 1,\ldots,n_1 \tag{6}$$

where $\mathbf{h}^{(1)} = \mathbf{p}^{(2)} - \mathbf{p}^{(1)}$ while $\nabla y_i^T(b,\mathbf{p}^{(1)})$ and $\mathbf{H}_i(b, \mathbf{p}^{(1)})$ signify the gradient and the hessian of the i-th function.

Relation (6) may also be expressed for all n_1 functions simultaneously in the following form:

$$\mathbf{y}(b,\mathbf{p}^{(2)}) \simeq \mathbf{y}(b,\mathbf{p}^{(1)}) + \mathbf{J}(b,\mathbf{p}^{(1)})\mathbf{h}^{(1)} + \frac{1}{2}\begin{bmatrix} \mathbf{h}^{(1)T}\mathbf{H}_1(b,\mathbf{p}^{(1)})\mathbf{h}^{(1)} \\ \mathbf{h}^{(1)T}\mathbf{H}_2(b,\mathbf{p}^{(1)})\mathbf{h}^{(1)} \\ \vdots \\ \mathbf{h}^{(1)T}\mathbf{H}_{n_1}(b,\mathbf{p}^{(1)})\mathbf{h}^{(1)} \end{bmatrix} \tag{7}$$

where $\mathbf{J}(b, \mathbf{p}^{(1)}))$ signifies the Jacobian matrix of function y_i, $i = 1, \ldots, n_1$.

Looking for vector $\mathbf{h}^{(1)}$, for which $\mathbf{y}(b, \mathbf{p}^{(2)}) = \mathbf{y}_b$, where $\mathbf{y}_b = [y_{1b}, y_{2b}, \ldots, y_{n_1 b}]^T$, we will obtain a non-linear set of equations for the coordinates of vector $\mathbf{p}^{(2)}$ in the form:

$$\mathbf{y}(b,\mathbf{p}^{(1)}) + \mathbf{J}(b,\mathbf{p}^{(1)})\mathbf{h}^{(1)} + \frac{1}{2}\begin{bmatrix} \mathbf{h}^{(1)T}\mathbf{H}_1(b,\mathbf{p}^{(1)})\mathbf{h}^{(1)} \\ \mathbf{h}^{(1)T}\mathbf{H}_2(b,\mathbf{p}^{(1)})\mathbf{h}^{(1)} \\ \vdots \\ \mathbf{h}^{(1)T}\mathbf{H}_{n_1}(b,\mathbf{p}^{(1)})\mathbf{h}^{(1)} \end{bmatrix} = \mathbf{y}_b \tag{8}$$

In the Newton-Raphson method based on the Taylor's formula including only the first derivatives, vector $\mathbf{h}^{(1)}$ is described with the following formula:

$$\mathbf{h}^{(1)} = -\mathbf{J}^{-1}(b, \mathbf{p}^{(1)})\,(\mathbf{y}(b, \mathbf{p}^{(1)}) - \mathbf{y}_b) \tag{9}$$

Substituting expression (9) into the third term of equation (8) we shall obtain

$$\mathbf{y}(b, \mathbf{p}^{(1)}) + \mathbf{J}(b, \mathbf{p}^{(1)})\mathbf{h}^{(1)} + \mathbf{r}(b, \mathbf{p}^{(1)}) = \mathbf{y}_b \tag{10}$$

where vector $\mathbf{r}(b, \mathbf{p}^{(1)})$ is built in the following way:

$$\mathbf{r}(b,\mathbf{p}^{(1)}) = \frac{1}{2}\begin{bmatrix} (\mathbf{y}(b,\mathbf{p}^{(1)})-\mathbf{y}_b)^T\mathbf{J}^{-T}(b,\mathbf{p}^{(1)})\mathbf{H}_1(b,\mathbf{p}^{(1)})\mathbf{J}^{-1}(b,\mathbf{p}^{(1)})(\mathbf{y}(b,\mathbf{p}^{(1)})-\mathbf{y}_b) \\ (\mathbf{y}(b,\mathbf{p}^{(1)})-\mathbf{y}_b)^T\mathbf{J}^{-T}(b,\mathbf{p}^{(1)})\mathbf{H}_2(b,\mathbf{p}^{(1)})\mathbf{J}^{-1}(b,\mathbf{p}^{(1)})(\mathbf{y}(b,\mathbf{p}^{(1)})-\mathbf{y}_b) \\ \vdots \\ (\mathbf{y}(b,\mathbf{p}^{(1)})-\mathbf{y}_b)^T\mathbf{J}^{-T}(b,\mathbf{p}^{(1)})\mathbf{H}_{n_1}(b,\mathbf{p}^{(1)})\mathbf{J}^{-1}(b,\mathbf{p}^{(1)})(\mathbf{y}(b,\mathbf{p}^{(1)})-\mathbf{y}_b) \end{bmatrix} \quad (11)$$

Solving equation (10) with respect to $\mathbf{h}^{(1)}$ we shall obtain the following iteration formula:

$$\mathbf{p}^{(k+1)} = \mathbf{p}^{(k)} - \mathbf{J}^{-1}(b,\mathbf{p}^{(k)})(\mathbf{y}(b,\mathbf{p}^{(k)}) - \mathbf{y}_b + \mathbf{r}(b,\mathbf{p}^{(k)})), \quad k = 1,2,\ldots \quad (12)$$

In order to be able to use formula (12), it is necessary to know the derivatives determining the Jacobian matrix $\mathbf{J}(b,\ \mathbf{p}^{(1)})$ and the Hessian matrices $\mathbf{H}_i(b,\ \mathbf{p}^{(1)})$, $i = 1,\ \ldots,\ n_1$. To this end we shall differentiate differential equations (1) with respect to individual parameters $p_j, j = 1, \ldots, n_1$.

$$\frac{\partial}{\partial p_j}\left(\frac{dy_i}{dx}\right) = \sum_{k=1}^{n} \frac{\partial f_i}{\partial y_k}\frac{\partial y_k}{\partial p_j}, \quad i = 1,\ldots,n, \quad j = 1,\ldots,n_1 \quad (13)$$

Defining the derivative sensitivity functions with the following formulas

$$g_{kj} = \frac{\partial y_k}{\partial p_j}, \quad k = 1,\ldots,n, \quad j = 1,\ldots,n_1 \quad (14)$$

we shall express equations (13) in the following form

$$\frac{dg_{ij}}{dx} = \sum_{k=1}^{n} \frac{\partial f_i}{\partial y_k} g_{kj}, \quad i = 1,\ldots,n, \quad j = 1,\ldots,n_1 \quad (15)$$

The initial conditions for the sensitivity function will be determined on the basis of formulas (3), (4) and (14)

$$g_{ij}(a) = \frac{\partial y_i(a)}{\partial p_j} = \delta_{ij}, \quad i = 1,\ldots,n, \quad j = 1,\ldots,n_1 \quad (16)$$

where δ_{ij} signifies the Kronecker symbol.

In order to determine the elements of the Hessian matrix, we shall differentiate differential equations (15) with respect to individual parameters p_r, $r = 1, \ldots, n_1$

$$\frac{\partial}{\partial p_r}\left(\frac{dg_{ij}}{dx}\right) = \sum_{k=1}^{n}\left[\sum_{q=1}^{n}\frac{\partial}{\partial y_q}\left(\frac{\partial f_i}{\partial y_k}\right)\frac{\partial y_q}{\partial p_r} g_{kj} + \frac{\partial f_i}{\partial y_k}\frac{\partial g_{kj}}{\partial p_r}\right], i = 1,\ldots,n,\ j = 1,\ldots n_1,\ r = 1,\ldots,n_1 \quad (17)$$

Defining the sensitivity functions derivatives with the following formulas

$$h_{jr}^{(k)} = \frac{\partial g_{kj}}{\partial p_r} = \frac{\partial^2 y_k}{\partial p_j \partial p_r}, \quad k = 1,\ldots,n, \quad j = 1,\ldots,n_1, r = 1,\ldots,n_1 \quad (18)$$

we shall express equations (17) in the following form

$$\frac{dh_{jr}^{(i)}}{dx}=\sum_{k=1}^{n}\left[\sum_{q=1}^{n}\frac{\partial}{\partial y_q}\left(\frac{\partial f_i}{\partial y_k}\right)g_{qr}g_{kj}+\frac{\partial f_i}{\partial y_k}h_{jr}^{(k)}\right],\quad i=1,\dots,n,\quad j=1,\dots,n_1,\quad r=1,\dots,n_1 \tag{19}$$

The initial conditions for the sensitivity function will be determined on the basis of formulas (16) and (18)

$$h_{jr}^{(i)}(a)=\frac{\partial g_{ij}(a)}{\partial p_r}=0,\quad i=1,\dots,n,\quad j=1,\dots,n_1,\quad r=1,\dots,n_1 \tag{20}$$

Finally, in each iteration we must solve the initial problem consisting of $n(1+n_1+n_1^2)$ equations (1), (15) and (19) and the same number of initial conditions determined by relations (3), (4), (16) and (20). Functions g_{ij} and $h_{jr}^{(i)}$ of indicator i less than or equal to n_1 compose the Jacobian matrix and the Hessian matrices respectively. The remaining functions appear only in differential equations (15) and (19).

4. Numerical example

Let us consider the boundary value problem in the following form

$$y''=\frac{1}{2}\frac{2(1+(y')^2)^{\frac{3}{2}}-(y')^2-1}{1.1-y},\qquad \begin{cases} y(0)=0\\ y'(1)=1\end{cases}$$

The above problem may be formulated with the use of a set of first-order equations in the form ($y_2=y$)

$$\begin{cases}\dfrac{dy_1}{dx}=\dfrac{1}{2}\dfrac{2(1+(y_1)^2)^{\frac{3}{2}}}{1.1-y_2}\\[2ex] \dfrac{dy_2}{dx}=y_1\end{cases},\qquad \begin{cases} y_1(1)=1\\ y_2(0)=0\end{cases}$$

In the problem under consideration we have: $n=2$, $n_1=1$. The initial problem that must be solved in each iteration consists of 6 first order differential equations

$$\frac{dy_1}{dx}=\frac{1}{2}\frac{2(1+(y_1)^2)^{\frac{3}{2}}-(y_1)^2-1}{1.1-y_2}$$

$$\frac{dy_2}{dx}=y_1$$

$$\frac{dg_{11}}{dx}=\frac{\partial f_1}{\partial y_1}g_{11}+\frac{\partial f_1}{\partial y_2}g_{21}$$

$$\frac{dg_{21}}{dx}=\frac{\partial f_2}{\partial y_1}g_{11}+\frac{\partial f_2}{\partial y_2}g_{21}$$

$$\frac{dh_{11}^{(1)}}{dx} = \frac{\partial^2 f_1}{\partial y_1^2} g_{11}^2 + 2\frac{\partial^2 f_1}{\partial y_1 \partial y_2} g_{21} g_{11} + \frac{\partial^2 f_1}{\partial y_2^2} g_{21}^2 + \frac{\partial f_1}{\partial y_1} h_{11}^{(1)} + \frac{\partial f_1}{\partial y_2} h_{11}^{(2)}$$

$$\frac{dh_{11}^{(2)}}{dx} = \frac{\partial^2 f_2}{\partial y_1^2} g_{11}^2 + 2\frac{\partial^2 f_2}{\partial y_1 \partial y_2} g_{21} g_{11} + \frac{\partial^2 f_2}{\partial y_2^2} g_{21}^2 + \frac{\partial f_2}{\partial y_1} h_{11}^{(2)} + \frac{\partial f_2}{\partial y_2} h_{11}^{(2)}$$

and 6 initial conditions

$$y_1(0) = p, \quad y_2(0) = 0, \quad g_{11}(0) = 1, \quad g_{21}(0) = 0, \quad h_{11}^{(1)}(0) = 0, \quad h_{11}^{(2)}(0) = 0$$

Iteration formula (12) will now adopt the form corresponding to one of the Halley's method variants [4]

$$p_{k+1} = p_k - \frac{2(y_1(b, p_k) - y_{1b}) g_{11}(b, p_k)^2 + (y_1(b, p_k) - y_{1b})^2 h_{11}^{(1)}(b, p_k)}{2 g_{11}(b, p_k)^3}, \quad k = 1, 2, \ldots$$

Adopting $h_{11}^{(1)}(b, p_k) = 0$ in the above iteration formula and taking into consideration the first 4 differential equations and 4 initial conditions we shall obtain Newton's approach characterized by quadratic convergence. The results of calculations for the quadratically convergent and cubically convergent methods have been presented in Table 1; in both approaches the parameter value $p_1 = 0$ has been adopted as the initial value.

Table 1

The convergence of presented method and Newton-Raphson method

Iterations	Parameter p values ($p_1 = 0$)	
	Newton-Raphson method	Cubically convergent method
1	0.**1**674150636	0.**1**029115260
2	0.**1**324421677	0.**115**7670195
3	0.**11**73361567	0.**1158044384**
4	0.**1158**168118	
5	0.**11580443**92	
6	0.**1158044384**	

As we can see, the number of significant figures in subsequent approximations triples if we use the presented method, while with the use of the Newton-Raphson approach it only doubles.

5. Conclusions

The method presented in the work requires that we know the Jacobian and Hessian matrices for functions with imposed boundary conditions at the final point of the domain. Replacing the boundary value problem with the initial value problem results in the increase of the number of differential equations from n to $n(1+n_1+n_1^2)$, with $n \cdot n_1$ of them pertaining to sensitivity function $g_{ij}(x)$ and $h_{jr}^{(i)}(x)$. pertaining to the derivatives of sensitivity function $h_{jr}^{(i)}(x)$. The definition of the above-mentioned matrices and formulation of the iteration formula and additional differential equations can be done in a simple way with the use of software for symbolic computation. The Newton-Raphson approach does not require knowledge of the Hessian matrix, hence the number of equations in the initial value problem is $n \cdot n_1^2$ less than in the approach presented in the work. The presented method may also be used to solve the boundary value problem in which some of the functions have boundary conditions imposed at both ends of the domain.

References

[1] Press W.H., Teukolsky S.A., Vetterling B.F., *Numerical Recipes. The Art of Scientific Computing*, third ed. Cambridge University Press, New York 2007.

[2] Palej R., *Cubically convergent method for nonlinear equation systems*, Czasopismo Techniczne 4-M/2011/B, Wydawnictwo PK, Kraków 2011.

[3] Rao S.S., *Applied Numerical Methods for Engineers and Scientists*, Prentice Hall, New Jersey 2002.

[4] Dahlquist G., Björck Å., *Numerical Methods in Scientific Computing*, Volume I, SIAM, Philadelphia 2008.